UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/810,232	03/26/2004	Ryan Rifkin	23085-07809	9459

758 7590 09/05/2007

FENWICK & WEST LLP
SILICON VALLEY CENTER
801 CALIFORNIA STREET
MOUNTAIN VIEW, CA 94041

EXAMINER	
RIDER, JUSTIN W	

ART UNIT	PAPER NUMBER
2626	

MAIL DATE	DELIVERY MODE
09/05/2007	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No.	Applicant(s)	
	10/810,232	RIFKIN, RYAN	
	Examiner	Art Unit	
	Justin W. Rider	2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *26 March 2004*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-44* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-44* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☒ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *36 March 200 and 04 August 2006* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

1. This action is responsive to communications: Application filed 36 March 2004. Claims 1-44 are pending.

Information Disclosure Statement

2. The information disclosure statement(s) (IDS) submitted on 10/2004 and 08/2006 are in compliance with the provisions of 37 CFR 1.97. Accordingly, the examiner has considered the information disclosure statement(s).

Specification

3. The abstract of the disclosure is objected to because it is too lengthy. Applicant is reminded that the abstract should be no longer than 150 words. Correction is required. See MPEP § 608.01(b).

Claim Rejections - 35 USC § 103

4. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

5. Claims 1,3-5, 8-9, 11-13, 15, 17, 19-21, 24-25, 27-29, 31, 33-35, 38-39 are rejected under 35 U.S.C. 103(a) as being unpatentable over **Chen et al. (US Patent No. 6,539,351)** referred to as **Chen** hereinafter, in view of **Pellom et al. ("an efficient scoring algorithm for Gaussian**

mixture model based speaker identification", IEEE signal processing letters, vo1.5, 1998, pages 28 1-284) referred to as **Pellom** hereinafter.

Claim 1: **Chen** discloses a method of voice enrollment and recognition, comprising:

i. organizing a plurality of speaker data points, representing a plurality of enrollment speakers, into a data structure using high dimensional vectors that represent characteristics of enrollment voice samples from the enrollment speakers (col. 2, lines 40-62, 'generating (organizing) a high dimensional density model (read on data structure)', 'transforming acoustic data (characteristics of enrollment voice samples --speaker data) obtained from at least one speaker (enrollment speaker) into high dimensional feature vectors (can also be interpreted as speaker data points);

ii. estimating a density of a subset of the plurality of speaker data points comprising the [approximate] nearest neighbors to an unidentified voice sample from an unidentified speaker (col. 1, lines 32-32 and 63-67, 'nearest neighbor methods', 'parametric density models' for 'density estimation', 'Gaussian mixture density models.. .with a relatively small number of parameters (corresponding to subset of the plurality of speaker data points)', wherein the estimation necessarily includes using unidentified voice sample from an unidentified speaker, so that it would have been obvious to one of ordinary skill in the art at time the invention was made to combine nearest neighbor methods and parametric density models, such as Gaussian mixture density model, for a high dimensional density estimation using relatively small number of parameters, as taught by **Chen** himself, for the purpose of offering 'decent performance with a relative small number of parameters' for the estimation); and

iii. identifying the unidentified speaker based on one or more speaker data points most closely matching the unidentified voice sample as indicated by the estimated density (col. 2, lines 40-62, 'provides.. . expectation maximization (most closely matching) method which estimates the parameters of the mixtures of the density model' for 'speaker reorganization (necessarily includes unidentified speaker and it voice sample).

Chen fails to explicitly disclose "approximate" nearest neighbors. However, the feature is well known in the art as evidenced by **Pellom** who discloses 'an efficient scoring algorithm for Gaussian mixture model based speaker identification (title), comprising 'approximated nearest-neighbor Gaussian mixture density evaluation' and 'nearest-neighbor approximation with beam-search' (page 283, left col., last paragraph).

Therefore, it would have been obvious to one of ordinary skill in the art at time the invention was made to modify **Chen** by providing an approximated nearest-neighbor and/or nearest-neighbor approximation for speaker identification as taught by **Pellom**, for the purpose of reducing the computational complexity of identifying a speaker (**Pellom**: abstract).

Claim 3: **Chen,** in view of **Pellom** discloses a method as per claim 1 above, wherein the step of estimating further comprises estimating the density based on a distance between individual speaker data points within the subset of speaker data points, (**Chen**: col. 13, lines 65-67, 'distance between their density functions'; col. 16, lines 27-29, 'minimizes the negentropy distance'; **Pellom**: page 282, right col., paragraph 2, 'form a subset of observations (speaker data points)').

Claim 4: **Chen,** in view of **Pellom** discloses a method as per claim 1 above, further comprising controlling the relative contributions of individual speaker data points within the

subset of speaker data points to the density based on a distance to a speaker data point from the unidentified voice sample, (**Pellom**: page 282, right col., paragraph 2, 'form a subset of observations (speaker data points) by sampling the observations nearest to the midpoint of previously scored element (interpreted as controlling the relative contributions)', 'pick highest probable speaker').

Claim 5: **Chen,** in view of **Pellom** discloses a method as per claim 1 above, further comprising estimating the density of the subset of speaker data points independent of parametric distribution information related to the plurality of speaker data points", (**Chen**: col. 2, lines 57-60, 'each compound Gaussian models each of the coordinated as each of the components of the density model in dependently'; **Pellom**: page 28 1, right mi., paragraph 1, 'in (3), the observations are assumed to be statistically independent').

Claim 8: **Chen,** in view of **Pellom** discloses a method as per claim 1 above, wherein the plurality of speaker data points comprises a relatively large number of speaker data points, (**Pellom**: page 283, left col., paragraph 1, '6 min of speech found in the enrollment section of database.. . ten tests per speaker each of approximately 15 s in duration.. . 10 ms from 20 ms overlapping window, 19 mel-frequency cepstral coefficients ... 64 Gaussian mixture...').

Claim 9: **Chen,** in view of **Pellom** discloses a method as per claim 1 above, further comprising retrieving the subset of speaker data points using an unidentified speaker data point from the unidentified voice sample as an index into the plurality of speaker data points", (**Pellom**: page 283, left col., paragraph 2, 'form a subset of observations (speaker data points) by sampling the observations nearest to the midpoint of previously scored element.. . spaced interval across the vectors', wherein i is index).

Claim 11: **Chen,** in view of **Pellom** discloses a method as per claim 1 above, further comprising more than one speaker data point associated with a common identification, and the step of identifying the unidentified speaker accumulates a score for the common identification", (**Pellom**: page 283, left col., paragraph 2, 'page 283, left col., paragraph 1, 'ten tests per speaker each of approximately 15 s in duration').

Claim 12: **Chen,** in view of **Pellom** discloses a method as per claim 1 above, further comprising identifying the unidentified speaker as one of the enrollment speakers if matching is within an error threshold", (**Chen** : col. 29, lines 2 1-45, 'extracting a lower dimensional feature from the original feature', 'estimated from the training data'; **Pellom**: page 282, right col., paragraph 1, steps 1)-5)).

Claim 13: **Chen,** in view of **Pellom** discloses a method as per claim 1 above, further comprising extracting the high-dimensional vectors from the enrollment voice samples and the unidentified voice sample", (**Chen** : col. 29, lines 50-58, 'speaker recognition'; **Pellom**: page 282, right col., paragraph 1, 'set a pruning threshold'(corresponding error threshold, 'pick (identifying) highest probable speaker').

Claim 15: Claim 15 is similar in scope and content to that of Claim 1 above and so is therefore rejected under the same rationale.

Claims 17, 19-21,24-25 and 27-29: Claims 17, 19-21, 24-25 and 27-29 are similar in scope and content to that of Claims 1, 3-5, 8-9 and 11-13 and so therefore are rejected under the same rationale.

Claims 31, 33, 35, 38-39 and 41-43: Claims 31, 33, 35, 38-39 and 41-43 are similar in scope and content to that of Claims 1, 3-5, 8-9 and 11-13 and so therefore are rejected under the same rationale.

6. Claims 2, 6, 16, 18, 22, 32 and 36 are rejected under 35 U.S.C. 103(a) as being unpatentable over **Chen**, in view of **Pollom** as applied to claim 1 above, and further in view of **Bahler (US Patent No. 5,271,088)** referred to as **Bahler** hereinafter.

Claim 2: **Chen,** in view of **Pellom** discloses a method as per claim 1 above, estimating a probability density function (**Chen**: col. 7, line 59). But, **Chen** in view of **Pellom** does not expressly disclose "using Patten windows to estimate the probability density function".

However, the feature is well known in the art as evidenced by **Bahler** who discloses automated sorting of voice messages through speaker sporting (title) for speaker recognition (abstract), comprising using Parzen estimate of local probability density and nearest neighbor distance (col. 8, lines 16-25).

Therefore, it would have been obvious to one of ordinary skill in the art at time the invention was made to modify **Chen** in view of **Pellom** by providing Parzen estimate for speaker identification, as taught by **Bahler**, for the purpose of improving speaker recognition system (**Bahler**: col. 5, lines 11-12).

Claim 6: **Chen,** in view of **Pellom** discloses a method as per claim 1 above, wherein a distance between individual speaker data points is based on characteristic similarities between associated voice samples, the distance measured in terms of one from the group containing: a Euclidean distance, a Minkowski distance, and a Manhattan distance.

However, the feature is well known in the art as evidenced by **Bahler** who further discloses 'the minimum Euclidean squared distance between the unknown speech frame and all reference frames of a given speaker over all frames of the unknown input (abstract).

Therefore, it would have been obvious to one of ordinary skill in the art at time the invention was made to modify **Chen** in view of **Pellom** by providing Euclidean squared distance, as taught by **Bahler**, for the purpose of improving speaker recognition (**Bahler**: abstract).

Claim 16, 18 and 32: Claims 16, 18 and 32 are similar in scope and content to that of Claim 2 above and so therefore is rejected under the same rationale.

Claims 22 and 36: Claims 22 and 36 are similar in scope and content to that of Claim 6 above and so therefore is rejected under the same rationale.

7. Claims 7, 10, 23, 26, 37 and 40 are rejected under 35 U.S.C. 103(a) as being unpatentable over **Chen**, in view of **Pollom** as applied to claims 1, 15 and 30 above, and further in view of **Arya et al., "an optional algorithm for approximated nearest neighbor searching in fixed dimensions", Journal of the ACM, Vol. 45, No.6, November 1998, pp.89 1-923)** referred to as **Arya** hereinafter.

Claim 7: **Chen**, in view of **Pollom** does not expressly disclose "the data structure comprises a kd-tree". However, the feature is well known in the art as evidenced by **Arya** who discloses 'an optional algorithm for approximated nearest neighbor searching in fixed dimensions' (title), comprising data structure 'implemented an optimized kd-tree' Page 914, paragraph 2).

Therefore, it would have been obvious to one of ordinary skill in the art at time the invention was made to modify **Chen** in view of **Pellom** by providing data structure using kd-tree, as taught by **Arya**, for the purpose of comparison of different approach for approximate nearest neighbor queries (**Arya**: page 9 14, paragraph 2).

Claim 10: **Chen**, in view of **Pollom** does not expressly disclose "retrieving approximate nearest neighbors to the unidentified speaker data point, the approximate nearest neighbors comprising speaker data points within a distance calculated as a function of a distance of an absolute nearest neighbor". However, the feature is well known in the art as evidenced by **Arya** who further discloses 'approximate nearest neighbor of q if its distance from q is within a factor of $(1+\varepsilon)$ of the distance to the true nearest neighbor (interpreted as absolute nearest neighbor)' (abstract).

Therefore, it would have been obvious to one of ordinary skill in the art at time the invention was made to modify **Chen** in view of **Pellom** by providing approximate nearest neighbor within a distance to the true nearest neighbor (absolute nearest neighbor), for the purpose of reducing computation time and/or complexity for data searching (**Arya**: abstract).

Claims 23 and 37: Claims 23 and 37 are similar in scope and content to that of Claim 7 above and so therefore is rejected under the same rationale.

Claims 26 and 40: Claims 26 and 40 are similar in scope and content to that of Claim 10 above and so therefore is rejected under the same rationale.

8. Claims 14, 30 and 44 are rejected under 35 U.S.C. 103(a) as being unpatentable over **Chen**, in view of **Pollom** as applied to claims 1, 15 and 30 above, and further in view of **Bahler et al. (US Patent No. 5,414,755)**, hereinafter referenced as **Bahler '755**.

Claim 14: **Chen,** in view of **Pellom** does not expressly disclose "an enrollment voice sample and the unidentified voice sample of a common speaker are text-independent".

However, the feature is well known in the art as evidenced by **Bahler '755** who discloses 'system and method for passive voice verification in a telephone network' (title), comprising 'a text-independent approach' to 'speaker recognition' (col. 4, lines 28-36).

Therefore, it would have been obvious to one of ordinary skill in the art at time the invention was made to modify **Chen** in view of **Pellom** by providing text-independent approach to speaker recognition, as taught by **Bahler '755**, for the purpose of improving speaker recognition system (**Bahler '755**: col. 5, lines 11-12).

Claims 30 and 44: Claims 30 and 44 are similar in scope and content to that of Claim 14 above and so therefore are rejected under the same rationale.

Conclusion

9. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. **Ashby et al. (US Patent No. 6,081,803)** discloses a system for system navigation utilizing groupings and kd-tree indexing.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Justin W. Rider whose telephone number is (571) 270-1068. The examiner can normally be reached on Monday - Friday 7:30AM - 5:00PM.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, David R. Hudspeth can be reached on (571) 272-7843. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

J.W.R.
27 August 2007

DAVID HUDSPETH
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2600